

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-342374
 (43)Date of publication of application : 29.11.2002

(51)Int.Cl. G06F 17/30
 G06T 1/00
 G06T 7/00

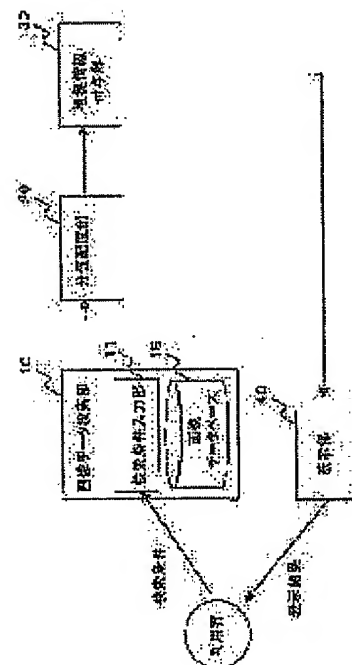
(21)Application number : 2001-143541 (71)Applicant : FUJITSU LTD
 (22)Date of filing : 14.05.2001 (72)Inventor : ENDO SUSUMU
 UEHARA YUSUKE
 MASUMOTO DAIKI

(54) SYSTEM AND METHOD FOR RETRIEVING DATA

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a data retrieving system, by which the contents of data retrieving processing can be easily grasped and the feature amount to be applied in re-retrieval can be easily selected.

SOLUTION: According to retrieval conditions inputted via a retrieval condition input part 11 by a user, a data acquiring part 10 retrieves data from a database 15. A classifying and arranging part 20 classifies and arranges the data of the retrieved result by using the feature amount. A base information applying part 30 applies base information with which the base of classifying and arranging the relevant data there is made visible, to each of data. For the base information, the distribution information of feature amounts in the relevant data is made visible, for example, and it is color histogram label information or wavelet conversion feature amount label information. According to the classification and arrangement information from the classifying and arranging part 20, a display part 40 displays each of data together with the corresponding base information.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-342374

(P2002-342374A)

(43) 公開日 平成14年11月29日 (2002. 11. 29)

(51) IntCl. ⁷	識別記号	F I	テームコード* (参考)
G 0 6 F 17/30	3 6 0	G 0 6 F 17/30	3 6 0 Z 5 B 0 5 0
	1 7 0		1 7 0 B 5 B 0 7 5
	3 5 0		3 5 0 C 5 L 0 9 6
G 0 6 T 1/00	2 0 0	G 0 6 T 1/00	2 0 0 E
7/00	3 0 0	7/00	3 0 0 F
審査請求 未請求 請求項の数10 O L (全 11 頁)			

(21) 出願番号 特願2001-143541(P2001-143541)

(22) 出願日 平成13年5月14日 (2001. 5. 14)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者 遠藤 進

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(72) 発明者 上原 祐介

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(74) 代理人 100095555

弁理士 池内 寛幸

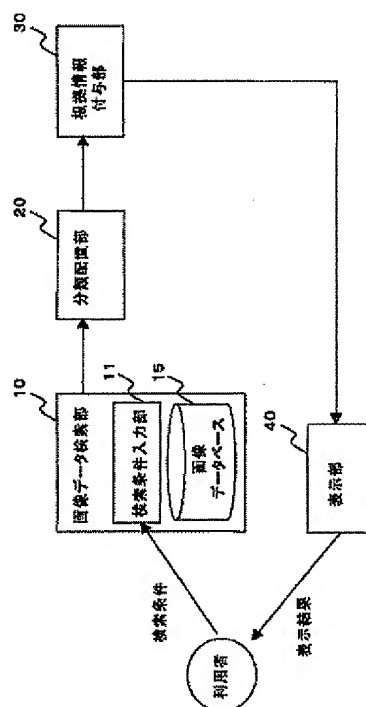
最終頁に続く

(54) 【発明の名称】 データ検索システムおよびデータ検索方法

(57) 【要約】

【課題】 データ検索処理の内容を容易に把握でき、再検索において与える特徴量の選定を容易としたデータ検索システムを提供する。

【解決手段】 検索条件入力部11を介して利用者が入力した検索条件に沿ってデータ取得部10はデータベース15からデータを検索する。分類配置部20は検索結果のデータを特徴量を用いて分類配置する。根拠情報付与部30は各データに対して、当該データがそこに分類配置された根拠を可視化した根拠情報を付す。根拠情報は例えば、当該データの持つ特徴量の分布情報を可視化したものであり、色ヒストグラムラベル情報やウェーブレット変換特徴量ラベル情報である。表示部40は、分類配置部20からの分類配置情報に従い、各データを対応する根拠情報とともに表示する。



【特許請求の範囲】

【請求項1】 データを取得するデータ取得部と、前記データ取得部のデータを特徴量に応じて表示画面上に分類配置する分類配置部と、各データに対し、前記分類配置部が当該データを分類配置した根拠を可視化した根拠情報を付す根拠情報付与部と、前記分類配置部からの分類配置情報に従い、各データに対応する根拠情報とともに表示する表示部とを備えたことを特徴とするデータ検索システム。

【請求項2】 前記表示部に表示されている根拠情報またはその一部分の情報、または、根拠情報の一部分を変更して得た情報を、再検索情報として受け付け、再検索処理を実行する請求項1に記載のデータ検索システム。

【請求項3】 前記分類配置部が、データの分類配置処理において、前記特徴量が近いデータ同士を近くに配置するように自己組織化マッピング処理を実行する請求項1または2に記載のデータ検索システム。

【請求項4】 前記根拠情報が、前記分類配置に用いた特徴量情報とそのデータが持つ前記特徴量の分布情報を示すラベル情報である請求項1に記載のデータ検索システム。

【請求項5】 前記データが画像データであり、前記ラベル情報において、前記特徴量情報が色情報であり、前記特徴量の分布情報が色ヒストグラム情報である請求項4に記載のデータ検索システム。

【請求項6】 前記データが画像データであり、前記ラベル情報において、前記特徴量情報がウェーブレット変換特徴量情報であり、前記特徴量の分布情報がウェーブレット変換された画像データの周波数成分の分布情報である請求項4に記載のデータ検索システム。

【請求項7】 前記データが画像データであり、前記ラベル情報において、前記特徴量情報が画像データ中におけるエッジ情報であり、前記特徴量の分布情報が画像データ中のエッジ部分の方向を要素としたヒストグラム情報である請求項4に記載のデータ検索システム。

【請求項8】 前記データが画像データであり、前記ラベル情報において、前記特徴量情報が色情報であり、前記特徴量の分布情報が代表色情報である請求項4に記載のデータ検索システム。

【請求項9】 データを取得し、前記データを特徴量に応じて表示画面上に分類配置し、各データに対し、その位置に分類配置した根拠を可視化した根拠情報を付し、各データを分類配置位置において、対応する根拠情報とともに表示することを特徴とするデータ検索方法。

【請求項10】 データ検索システムを実現するコンピュータプログラムであって、データを取得するデータ取得処理ステップと、前記データ取得処理ステップにより得られたデータを特

微量に応じて表示画面上に分類配置する分類配置処理ステップと、

各データに対し、前記分類配置処理ステップにおいて、その位置に分類配置した根拠を可視化した根拠情報を付す根拠情報付与処理ステップと、

各データを分類配置位置において、対応する根拠情報とともに表示する表示処理ステップとを備えたことを特徴とするコンピュータプログラム。

【発明の詳細な説明】

10 【0001】

【発明の属する技術分野】本発明は、大量のデータの中から目的のデータを柔軟かつ効率良く検索するデータ検索システムおよび方法に関する。

【0002】

【従来の技術】従来、大量のデータの中から目的とするデータを検索する方法としては、データを説明するテキスト情報を検索キーとして利用して検索処理を実行する方法と、データ中の特徴量を検索キーとして利用して検索処理を実行する方法が知られている。

20 【0003】前者のテキスト情報を検索キーとして利用する方法に関しては、第1に検索精度に問題があった。つまり、データを的確に示すテキスト情報を付与することは容易ではないため、データに付与されたテキスト情報が不十分である場合が多い。この場合、利用者が指定したテキスト情報を検索キーとした検索処理では目的とするデータの検索に手間取り、検索精度が向上しないという問題があった。第2に前処理に時間がかかるという問題があった。つまり、各データに当該データを的確に表わすテキスト情報を自動付与することは一般には困難であり、現状、自動付与が適正であるか人手による確認を必要としたり、人手によりテキスト情報を追加付与するという作業を行わざるを得ない場合が多く、前処理に相当の時間がかかっているのが現状である。

30 【0004】そこで、後者の、データ中の特徴量を検索キーとして利用して検索処理を実行する方法が注目されている。従来技術における、特徴量を検索キーとして利用した検索処理では、検索に用いる特徴量に注目し、各データが持つ特徴量を評価することにより、特徴量が類似したデータを見つけ出す方法が開発されている。

40 【0005】

【発明が解決しようとする課題】しかし、従来技術における、特徴量を検索キーとして利用するデータ検索処理には以下の問題があった。

【0006】人間の主観も加え、人間の認識の感覚を数値化することは困難であり、目的のデータを特定するためにどのような特徴量を用いても、人間の感覚にぴたり合う特徴量はないため、目的とするデータを特定するための最適な特徴量を定めることは困難である。そのため、試行錯誤的に何らかの特徴量を検索キーとして与え、データ検索を取り敢えず実行することが多い。この

場合、検索結果の中に、なぜそのようなデータが検索結果として表示されているのかが理解しづらい場合も多い。テキスト情報を検索キーとして利用する検索の場合は、データの中に当該テキスト情報をキーワードとする要素が含まれているかどうかを判断しやすく、ある程度検索結果は理解し易い。

【0007】また、試行錯誤的にデータ検索を続行して行くにあたり、テキスト情報を検索キーとして利用する検索の場合は、用いられたテキスト情報とその検索結果に基づいて、次の検索キーとして与えるテキスト情報を想起することも比較的容易である。しかし、特徴量を検索キーとして利用する方法においては、上記のごとく、再検索においても目的とするデータを特定するための最適な特徴量を定めることは困難である。どのような特徴量を与えて試行すれば良いのか経験的に把握することは難しい。

【0008】本発明は、特徴量を検索キーとして利用するデータ検索処理において、検索結果となったデータが抽出された根拠を分かりやすく表示し、検索処理の内容を概ね把握でき、再検索において与える特徴量の選定を容易としたデータ検索システム、データ検索方法およびデータ検索システムを実現する処理プログラムを提供することを目的とする。

【0009】

【課題を解決するための手段】上記目的を達成するために本発明のデータ検索システムは、データを取得するデータ取得部と、前記データ取得部のデータを特徴量に応じて表示画面上に分類配置する分類配置部と、各データに対し、前記分類配置部が当該データを分類配置した根拠を可視化した根拠情報を付す根拠情報付与部と、前記分類配置部からの分類配置情報に従い、各データに対応する根拠情報とともに表示する表示部とを備えたことを特徴とする。

【0010】上記構成により、検索結果として分類配置された各データに、分類配置の根拠情報が併せて表示されることとなり、根拠情報を見れば、そのデータの特徴量がどのようなものであり、データ検索処理内容が如何なるものであったかが把握しやすく、また、データ検索処理の結果、なぜ当該データがその位置に配置されているのかが一目で把握しやすくなる。

【0011】次に、上記構成において、前記表示部に表示されている根拠情報またはその一部分の情報、または、根拠情報の一部分を変更して得た情報を、再検索情報として受け付け、再検索処理を実行することが好ましい。

【0012】上記構成により、再検索を実行する際、データ検索結果を踏まえ、根拠情報として表示されている情報、例えば、色ヒストグラム情報中の特定の色情報、特定の周波数成分、根拠情報の一部分を変更して得た情報などを、後続のデータ検索におけるデータ検索キーと

して用いることができる。

【0013】次に、上記構成において、前記分類配置部が、データの分類配置処理において、前記特徴量が近いデータ同士を近くに配置するように自己組織化マッピング処理を実行することが好ましい。

【0014】上記構成により、データ検索の結果、データ検索に用いた特徴量が近いもの同士が近くに配置されることとなり、検索結果の一覧性が向上し、検索結果の内容を把握しやすくなる。

【0015】次に、上記構成において、前記根拠情報が、前記分類配置に用いた特徴量情報とそのデータが持つ前記特徴量の分布情報を示すラベル情報であることが好ましい。

【0016】上記構成により、検索結果として分類配置された各データにラベル情報が併せて表示されることとなり、ラベル情報を見れば、そのデータの特徴量がどのようなものであり、データ検索処理内容が如何なるものであったかが把握しやすく、また、データ検索処理の結果、なぜ当該データがその位置に配置されているのかが容易に把握できる。

【0017】なお、前記データが画像データであり、前記ラベル情報の特徴量情報を色情報とし、特徴量の分布情報を色ヒストグラム情報とすれば、画像データ検索の結果、選ばれた画像データの色に関する情報を俯瞰することができ、画像データ検索処理の内容を把握しやすく、また、色ヒストグラムの要素比を調整するなど、再検索に用いる特徴量を与える手掛かりが得られる。

【0018】また、前記データが画像データであり、前記ラベル情報の特徴量情報をウェーブレット変換特徴量情報とし、特徴量の分布情報を周波数成分の分布情報とすれば、画像データ検索の結果、選ばれた画像データの画像の周波数成分の情報を俯瞰することができ、画像データ検索処理の内容を把握しやすく、また、特定の周波数成分に注目するなど、再検索に用いる特徴量を与える手掛かりが得られる。

【0019】また、前記データが画像データであり、前記ラベル情報の特徴量情報を画像データ中におけるエッジ情報とし、特徴量の分布情報を画像データ中のエッジ部分の方向を要素としたヒストグラム情報とすれば、画像データ中に含まれるエッジが、縦、横、斜めの8方向のどちら方向を向いているものが多いのかという情報を表示することが可能となる。

【0020】また、前記データが画像データであり、前記ラベル情報の特徴量情報を色情報とし、特徴量の分布情報を代表色情報とすれば、画像データが含む概ねの色情報の傾向を知ることができ、画像データの色合いを手掛かりとして画像データ検索処理を実行することができる。

【0021】なお、本発明のデータ検索システムを実現する処理ステップを備えた処理プログラムコードを提供

10

20

30

40

50

すれば、コンピュータを用いて本発明の利用者認証システムを構築することができる。

【0022】

【発明の実施の形態】本発明のデータ検索システムについて図面を参照しつつ説明する。なお、以下の実施形態において、データの一例として画像データを挙げて説明しているが、データは画像データに限らず、音声データやテキストデータなど特徴量を抽出し得るものであれば良い。

【0023】（実施形態1）実施形態1のデータ検索システムは、データを取得し、そのデータの特徴量に応じて表示画面上に分類配置し、表示の際に各データに対して当該位置に分類配置された根拠情報を付し、データの分類配置結果を表示するものである。

【0024】図1が本発明のデータ検索システムの概略構成を示す図である。

【0025】10はデータ取得部である。分類配置対象となるデータをあらかじめ取得・格納しておくものでも良く、また、データ検索処理能力を持ち、検索条件に沿ってデータベースからデータを検索・格納するものでも良い。ここでは後者のタイプとして説明する。データ取得部10は、データ検索処理能力を持ち、データベース15にアクセスし、所定の検索アルゴリズムに従って、検索条件に沿ってデータを検索・取得する。なお、データベース15は、本発明のデータ検索システム内に備えられたものでも良く、外部のデータベース、例えば、インターネットを介して接続された外部のデータベースであっても良い。図1では、データ検索システム内に備えられている例を示している。なお、外部のデータベースを利用する場合、データ取得部10は、外部のデータベースを管理するデータ検索エンジンに対して、検索条件を示してデータの検索を依頼し、検索されたデータを受け取るものでも良い。

【0026】検索条件は検索条件入力部11を介して利用者から入力され、検索条件としては、特徴量を検索キーとして与えるもので良く、さらに、テキスト情報を検索キーとして併用するものでも良い。

【0027】20は分類配置部であり、データ取得部10に取得されているデータをその特徴量に応じて表示画面上に分類配置する。本発明のデータ検索システムにおいては、分類配置部20が用いる分類配置のアルゴリズムは様々なものを適用することができる。一例として、自己組織化マッピングアルゴリズムを用いることができる。

【0028】自己組織化マップ技術は、ニューロ技術を応用した分類技術の一つで、局所相互作用を利用して、学習により自己組織化をおこなうものである。この手法をデータの分類配置に応用すれば、マッピング空間上でデータの注目している特徴量が近いと思われるものが近くなるように配置することが出来る。このように、注

目している特徴量が近いデータ同士が近くに配置されることにより、データ検索システム利用者はデータ配置を一覧することが容易となり、データ分類配置の様子が分かりやすくなる。

【0029】30は根拠情報付与部であり、それぞれのデータに対して当該データが所定位置に分類配置された根拠を可視化した根拠情報を付す。

【0030】根拠情報は、データ検索システムの利用者に対して、当該データがそこに分類配置された根拠を端的に分かりやすく可視化した情報であることが好ましい。根拠情報の一例は、分類配置に用いた特徴量情報とそのデータが持つ特徴量の分布情報を示すラベル情報である。さらにラベル情報として様々なラベル情報が考えられる。

【0031】ラベル情報の一例は、データが画像データであり、特徴量情報が色情報であり、特徴量の分布情報が色ヒストグラム情報としたものである。図2は、データを画像データとし、根拠情報として、特徴量情報が色情報であり、特徴量の分布情報が色ヒストグラム情報としたラベル情報を画像データに付した例を示す図である。図2に見るように、画像データ201に対してキャプションのように色ヒストグラムラベル情報202が付されている。この例では簡単に画像データ201は3つの色領域部分を持ち、図示したような比率の色ヒストグラムを持っている。このように、各画像データに色ヒストグラムラベル情報202を付しておくことにより、画像データ検索システム利用者は、画像データ中の色情報が画像データ検索および画像データ分類配置結果に与える影響を容易に把握できる。

【0032】ラベル情報の他の例は、データが画像データであり、特徴量情報がウェーブレット変換特徴量情報であり、特徴量の分布情報が周波数成分の分布情報としたものである。ウェーブレット変換を利用すれば、画像データにおいて、縦、横、斜めのそれぞれの方向に関して、周波数成分を抽出することができる。図3は、根拠情報として、特徴量情報がウェーブレット変換特徴量情報であり、特徴量の分布情報が周波数成分の分布情報としたラベル情報を画像データに付した例を示す図である。図3に見るように、画像データ201に対してキャプションのようにウェーブレット変換特徴量ラベル情報203が付されている。

【0033】図3に示したウェーブレット変換特徴量ラベル情報203は3つの部分に分かれている。それぞれ、縦方向の周波数分布情報203a、横方向の周波数分布情報203b、斜め方向の周波数分布情報203cである。これら3つの周波数分布情報203a～cは、周波数を横軸としており、左側が低周波成分、右側が高周波成分をあらわしている。ウェーブレット変換により周波数変換した場合、周波数成分情報とともに位置情報を保持できるため、例えば低周波成分がどこにどれくら

10

20

30

40

50

い存在しているかなどといった情報を調べることが出来る。このように、各画像データにウェーブレット変換特徴量ラベル情報203を付しておくことにより、データ検索システム利用者は、画像データ中の縦、横、斜め方向における周波数成分、つまり、画像部分の縦、横、斜めの各方向における形状、変化の細かさなどの情報が、画像データ検索および画像データ分類配置結果に与える影響を容易に把握できる。

【0034】ラベル情報の他の例は、データが画像データであり、特徴量情報が色情報であり、特徴量の分布情報が代表色情報とした代表色ラベル情報である。代表色を一色のみとしても良いし、2色、3色など複数としても良い。それら代表色の色ヒストグラムをラベル情報としても良い。代表色が一色のみの場合は、特徴量の分布情報が当該色1つのみとなり、代表色が複数の場合は、特徴量の分布情報が当該複数色の比を示すこととなる。ここで、代表色とは、画像データを代表する色である。根拠情報付与部30が代表色を決めるアルゴリズムは多様なものを採用することが可能である。例えば、単純に色領域の大きさが抽出できる場合には、最も領域の大きな色を代表色とすることができる。

【0035】また、画像データ中、色領域がディザ状になっている場合などでは画像データ中の色集約度などを算出し、色領域の塊を見つけるというアルゴリズムを採用することができる。また、このように根拠情報付与部30が選定した代表色を手で確認し、代表色を選定し直すことも可能である。このように、代表色ラベル情報を画像データに付しておくことにより、データ検索システム利用者は、画像データが含む概ねの色情報の傾向を知ることができ、画像データの色合いを手掛かりとして画像データ検索処理を実行することができる。

【0036】ラベル情報の他の例は、データが画像データであり、特徴量情報が、画像データ中におけるエッジ情報であり、特徴量の分布情報が、エッジ部分の方向を要素としたヒストグラム情報としたものである。このエッジヒストグラムラベル情報によって、画像データ中に含まれるエッジが、縦、横、斜めの8方向のどちら方向を向いているものが多いのかという情報を表示することが可能となる。

【0037】40は表示部であり、分類配置部20から得られた分類配置情報に従い、各データに対応する根拠情報とともに表示する。

【0038】なお、図1のデータ検索システムの概略構成において、通信インタフェースやその他のデバイスなどは説明の便宜上、図示を省略している。

【0039】次に、本発明のデータ検索システムを用いたデータ検索の処理の流れを示す。なお、データとして画像データを例に挙げて説明する。

【0040】図4は、本発明のデータ検索システムを用いたデータ検索処理の流れの一例を示すフローチャート

である。

【0041】なお、以下の例は、データ取得部10が、データ検索処理能力を持ち、データの取得にあたり、データの検索処理が実行される例である。分類配置対象となるデータがあらかじめデータ取得部10内に取得されている場合には、かかる処理ステップは必要ない。

【0042】まず、前処理として、データベース15に格納する画像データから所定の特徴量に関する分布情報を抽出しておく。これらは後述するデータの検索処理、分類配置処理において用いるものであり、例えば、色ヒストグラム情報、ウェーブレット変換特徴量情報、代表色情報、エッジヒストグラム情報などである。

【0043】利用者は、データ取得部10に対して、検索条件入力部11を介して画像データの検索条件を入力する(ステップS401)。検索条件としては、所望の画像の特徴の内容を示すキーワードや、利用したい特徴量の種別とそのパラメータなど多様なものがあり得る。ここでは例えば、利用したい特徴量の種別として色情報を指定し、“青色”をパラメータとして与える。

【0044】次に、データ取得部10は、入力された検索条件に従い、画像データをデータベース15から検索する(ステップS402)。データ取得部10は、検索結果として得られた画像データを分類配置部20に渡す。

【0045】次に、分類配置部20は、データ取得部10から渡された画像データを分類配置する(ステップS403)。分類配置は、1次元に並べるだけでなく、自己組織化マップなどの手法を用いて、2次元、3次元に分類配置することも可能である。なお、このステップS403の分類配置の特徴は、ステップS401で検索条件として用いた特徴量やパラメータなどを分類配置の際の特徴量として用いることも可能であり、また、ステップS401で検索条件として用いた特徴量やパラメータなどとは別の特徴量を用いることも可能である。前者の場合であれば、分類配置部20は、データ取得部10の検索条件入力部11から入力された検索条件として用いた特徴量やパラメータなどを受け取り、分類配置用の特徴量として用いる。後者の場合であれば、このステップS403において、利用者などから分類配置用の特徴量の指定を受け付けるか、あらかじめ分類配置用の特徴量を決めておく。

【0046】次に、根拠情報付与部30は、各画像データごとに検索処理、分類配置処理に用いた特徴量を基に根拠情報を作成して画像データに付与する(ステップS404)。例えば、特徴量として色情報が用いられている場合には、色ヒストグラムラベル情報を作成して各画像データに付与する。

【0047】次に、表示部40は、画像データと当該画像データに付与された根拠情報とを併せて分類配置位置に並べて表示する(ステップS405)。

10

20

30

40

50

【0048】図5は、分類配置された画像データの表示例を示す図である。この例は、画像データに色ヒストグラムラベル情報が根拠情報として付与されている場合の分類配置例である。画像データは簡単に9つのみを模式化して示している。図5に示すように各画像データの下側には色ヒストグラム情報を表わすラベル情報が付与されている。データ検索システム利用者は、このラベル情報を参考にするにより、各画像データの検索結果、分類配置結果において各色領域の割り合いが与える影響を概ね把握することができる。

【0049】図6は、分類配置された画像データの別の表示例を示す図である。この例は、画像データに対してウェーブレット変換周波数成分の分布情報を表わすウェーブレット変換特徴量ラベル情報が根拠情報として付与されている場合の分類配置例である。画像データは簡単に9つのみを模式化して示している。図6に示すように各画像データの下側にはウェーブレット変換特徴量ラベル情報が付与されている。ウェーブレット変換特徴量ラベル情報中の3つの区分のうち、左側が縦方向の周波数成分分布、中側が横方向の周波数成分分布、右側が斜め方向の周波数成分分布を示している。データ検索システム利用者は、このラベル情報を参考にするにより、各画像データの検索結果、分類配置結果において各方向の周波数成分分布が与える影響を概ね把握することができる。

【0050】以上、本実施形態1に示すデータ検索システムは、検索結果として分類配置された各画像データに、分類配置の根拠情報が併せて表示することができ、利用者は、色ヒストグラムラベル情報、ウェーブレット変換特徴量ラベル情報、代表色ラベル情報、エッジヒストグラムラベル情報などの根拠情報を見れば、その画像データの特徴量がどのようなものであり、画像データ検索処理内容が如何なるものであったか、また、どのような基準で当該画像データがその位置に配置されているのかが容易に把握できる。

【0051】（実施形態2）実施形態2にかかるデータ検索システムは、実施形態1に示したデータ検索システムをさらに改良したものであり、表示部に表示されたデータを再検索する場合の工夫を盛り込んだものであり、表示部に表示されている根拠情報またはその一部分の情報、または、根拠情報の一部分を変更して得た情報を、再検索情報として受け付け、再検索処理を実行するものである。

【0052】なお、本実施形態2においても、データ取得部10がデータ検索処理能力を持ち、データベース15からデータ検索を行ってデータを取得する場合の構成を挙げる。

【0053】実施形態2のデータ検索システムの概略構成図は、図1と同様で良いが、検索条件入力部11における処理が異なる。実施形態2の検索条件入力部11

は、後述する再検索条件の入力が可能なものである。実施形態2の検索条件入力部11は、データに付された根拠情報を利用した簡単かつ柔軟な再検索条件の入力ができるものである。

【0054】例えば、実施形態1と同様、データを画像データとし、図2に示した色ヒストグラムラベル情報を利用した再検索条件入力の様子を図7に示す。図7に示すように、表示画面中のある画像データの色ヒストグラムラベル情報のある色要素をマウスなどのポインティングデバイスのカーソル210でクリックすると、検索条件入力部11は、再検索条件として当該色情報がパラメータとして指定されたと認識する。なお、併せて再検索の処理区分の入力も可能とする。

【0055】「検索やり直し」の処理区分を指定して、再検索条件として色情報を与えれば、当該色情報を検索キーとして、データ取得部10がデータベース15に対して再度アクセスを行い、画像データを再検索する。例えば、再検索前において「青」の色情報が検索キーとして分類配置されており、「検索やり直し」において「赤」の色情報が検索キーとして与えられれば、データベース15から「赤」を代表色として持つ画像データなどが再検索・再取得され、分類配置部20により分類配置されることとなる。

【0056】「検索絞り込み」の処理区分を指定して、再検索条件として色情報を与えれば、現時点で、データ取得部10に得られている画像データに対して、さらに当該色情報を検索キーとして絞り込み検索を行う。例えば、再検索前において「青」の色情報が検索キーとして分類配置されており、「検索絞り込み」において「赤」の色情報が検索キーとして与えられれば、分類配置されている画像データのうち、「青」および「赤」の両方の色情報を多く含む画像データが絞り込まれることとなる。

【0057】「再分類配置」を指定して再検索条件として色情報を与えれば、現在得られている画像データを基に、当該色情報をパラメータとして再度、分類配置をやり直す。再分類配置の処理方法には多数の方法が想定できる。第1の処理方法は、再分類配置に指定された情報以外の情報を無視して、分類配置そのものをやり直す方法である。例えば、再分類配置前において「青」の色情報を基に分類配置されており、第1の処理方法を用いて「赤」の色情報を基に再分類配置を実行すると、再分類配置に用いる画像データは再分類配置前の画像データと同じであるが、再分類配置では「青」の色情報は無視され、「赤」の色情報を基にした新しい分類配置が行われる。第2の処理方法は、再分類配置前の分類配置の結果を引き継いで、再分類配置に指定された情報を加えて分類配置する方法である。例えば、再分類配置前において「青」の色情報を基に分類配置されており、第2の処理方法を用いて「赤」の色情報を基に再分類配置を実行す

ると、再分類配置に用いる画像データは再分類配置前の画像データと同じであり、再分類配置前の「青」の色情報に基づく分類配置結果について、さらに「赤」の色情報を基にした分類配置を実行する。この場合、「青」の色情報に基づく分類配置結果は全体としてあまり変化しないが、その中で「赤」の色情報を多く持つ画像データが一まとめに分類配置される。

【0058】なお、オプションとして、再分類配置前の分類配置画面において特定の画像を指定して位置を固定して再分類配置後、同じ位置に配置されるように指定したり、再分類配置後は中央に配置されるように指定したりすることができる。

【0059】再分類配置の一例を図で示す。例えば、ある検索条件による検索処理が実行され、分類配置結果として、図5の色ヒストグラムラベル情報を持つ画像データが表示されているとする。ここで、利用者は図5の分類配置結果を見て、これら画像データのうち、今度は赤い色（図5、図8では赤色領域を縦線のハッチングで示している。）に注目して再分類配置したいと考えたとする。ここで、マウスなどのポインティングデバイスのカーソル210で、色ヒストグラムラベル情報中の赤色要素をクリックし、上記第1の再分類配置処理方法による「再分類配置」を指定すると、図5に示す画像データの分類配置が、図8に示すように再分類配置される。

【0060】なお、図9に示すように、表示画面中のある画像データの色ヒストグラムラベル情報の色要素間の境界線をドラッグして移動させることにより、色ヒストグラム情報を変更して、再検索条件として指定することも可能である。この場合、再検索条件として指定された色ヒストグラム情報を持つ画像データ、当該色ヒストグラム情報に対するマッチング度合いが高い所定数の画像データが検索される。

【0061】他の例を示す。図3に示したウェーブレット変換特徴量ラベル情報を利用した再検索条件入力の様子を図10に示す。図10に示すように、表示画面中のある画像データのウェーブレット変換特徴量ラベル情報のある方向の周波数成分分布をマウスなどのポインティングデバイスのカーソル210でクリックすると、検索条件入力部11は、再検索条件として当該方向に関する周波数成分分布がパラメータとして指定されたと認識する。併せて再検索の処理区分の入力も可能であり、「検索やり直し」の処理区分や「検索絞り込み」の処理区分や「再分類配置」の処理区分の指定が可能である。今、ある検索条件による検索処理が実行され、分類配置結果が表示されているとする。例えば、図6のウェーブレット変換特徴量ラベル情報を持つ画像データが表示されているとする。例えば、利用者は図6の分類配置結果を見て、これら画像データのうち、今度は縦方向の周波数成分に注目して再分類配置したいと考えたとする。こ

こで、マウスなどのポインティングデバイスのカーソル2

10で、ウェーブレット変換特徴量ラベル情報中の縦方向の周波数成分をクリックし、「再分類配置」を指定すると、図6に示す画像データの分類配置が、図11に示すように再分類配置される。

【0062】また、図12に示すように、表示画面中のある画像データのウェーブレット変換特徴量ラベル情報の周波数成分の包絡線をドラッグして変形することにより、ウェーブレット変換特徴量ラベル情報を変更して、再検索条件として指定することも可能である。この場合、再検索条件として指定されたウェーブレット変換周波数成分を持つ画像データ、当該ウェーブレット変換周波数成分に対するマッチング度合いが高い所定数の画像データが検索される。

【0063】このように、画像データの表示結果を見て、根拠情報により現在の検索処理、分類配置処理の内容を把握し、さらに、ラベル情報を用いて次の検索条件を適切かつ柔軟に指定することができる。

【0064】（実施形態3）本発明のデータ検索システムは、上記に説明した構成を実現する処理ステップを記述したプログラムを提供することにより、各種コンピュータを用いて構築することができる。本発明のデータ検索システムを実現する処理ステップを備えたプログラムをコンピュータにロードする方法は、当該プログラムを記録した記録媒体をコンピュータに読み取らせることでも良く、また、ネットワーク上からダウンロードすることでも良い。例えば、図13に図示したように、CD-ROM302やフレキシブルディスク303等の可搬型記録媒体301をコンピュータ304に読み取らせたり、ネットワーク上にある記録装置内の記録媒体300からネットワークを介してダウンロードすることができる。また、当初からコンピュータ304内のハードディスクやRAM等の記録媒体305に格納していても良い。当該データ検索システムを実現する処理プログラムは、プログラム実行時にコンピュータ304上にローディングされ、主メモリ上で実行される。

【0065】なお、本発明のデータ検索システムは、上記の発明の概念から逸脱することなく、上記の方法及び装置に種々の変更及び変形を成し得ることが理解されよう。従って、本発明は上記実施形態に限定されるものではないことに注意する必要がある。

【0066】

【発明の効果】本発明のデータ検索システムによれば、検索結果として分類配置された各データに、分類配置の根拠情報が併せて表示されることとなり、利用者が当該根拠情報を見れば、そのデータの特徴量がどのようなものであり、データ検索処理内容が如何なるものであったかが把握しやすく、また、データ検索処理の結果、なぜ当該データがその位置に配置されているのかが容易に把握できる。

【0067】また、本発明のデータ検索システムによれ

ば、再検索を実行する際、データ検索結果を踏まえ、データ検索で用いた特徴量の特定の要素、例えば、特定の色情報、特定の周波数成分などを直接、後続のデータ検索におけるデータ検索キーとして用いたり、データ検索に用いたい根拠情報または根拠情報の一部の情報を指定し、または根拠情報の一部の情報を変更した上で再データ検索における検索キーとして用いることができ、適切かつ柔軟に再検索条件の設定が可能となる。

【図面の簡単な説明】

【図1】 本発明の実施形態1のデータ検索システムの概略構成を示す図

【図2】 根拠情報として、特徴量情報が色情報であり、特徴量の分布情報が色ヒストグラム情報としたラベル情報を画像データに付した例を示す図

【図3】 根拠情報として、特徴量情報がウェーブレット変換特徴量情報であり、特徴量の分布情報が周波数成分の分布情報としたラベル情報を画像データに付した例を示す図

【図4】 本発明のデータ検索システムを用いたデータ検索処理の流れの一例を示すフローチャート

【図5】 本発明のデータ検索システムにより分類配置された画像データの表示例を示す図

【図6】 本発明のデータ検索システムにより分類配置された画像データの別の表示例を示す図

【図7】 図2に示した色ヒストグラムラベル情報を利用した再検索条件入力の様子を示す図

【図8】 図5に示す画像データの分類配置が再分類配置される様子を示す図

【図9】 図2に示した色ヒストグラムラベル情報の色要素間の境界線をドラッグして移動させて色ヒストグラム情報を変更し、再検索条件として指定する様子を示す図

【図10】 図3に示したウェーブレット変換特徴量ラ*

* ベル情報を利用した再検索条件入力の様子を示す図

【図11】 図6に示す画像データの分類配置が再分類配置される様子を示す図

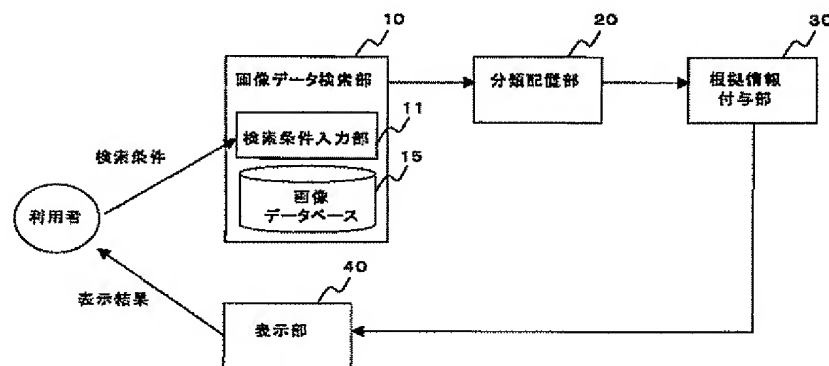
【図12】 図3に示したウェーブレット変換特徴量ラベル情報の周波数成分の包絡線をドラッグして変形し、再検索条件として指定する様子を示す図

【図13】 本発明のデータ検索システムを実現する処理ステップのプログラムをコンピュータにロードする様子を示す図

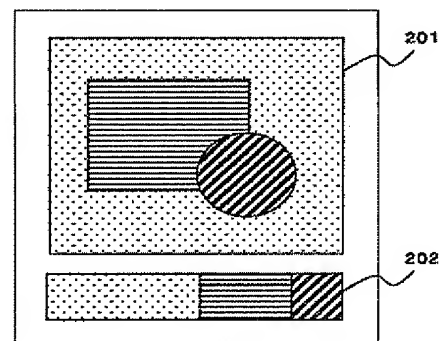
【符号の説明】

- 10 データ取得部
- 11 検索条件入力部
- 15 データベース
- 20 分類配置部
- 30 根拠情報付与部
- 40 表示部
- 201 画像データ
- 202 色ヒストグラムラベル情報
- 203 ウェーブレット変換特徴量ラベル情報
- 203a ウェーブレット変換の縦方向の周波数分布情報
- 203b ウェーブレット変換の横方向の周波数分布情報
- 203c ウェーブレット変換の斜め方向の周波数分布情報
- 300 ネットワーク上にある記録装置内の記録媒体
- 301 可搬型記録媒体
- 302 CD-ROM
- 303 フレキシブルディスク
- 304 コンピュータ
- 305 コンピュータのハードディスクやRAM等の記録媒体

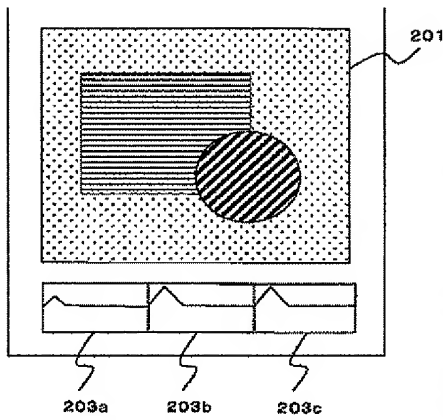
【図1】



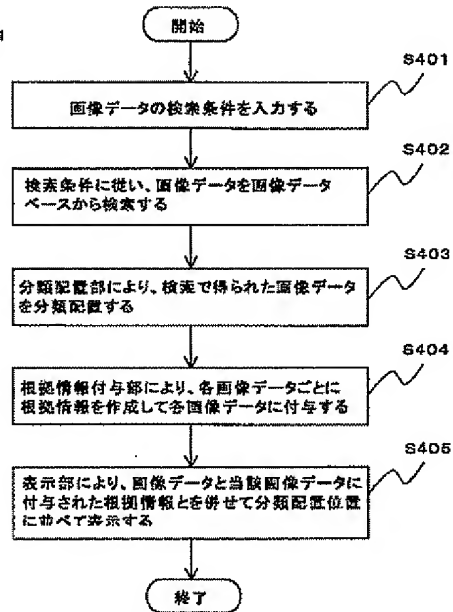
【図2】



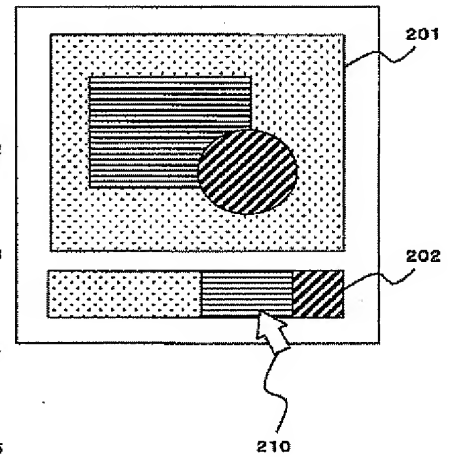
【図3】



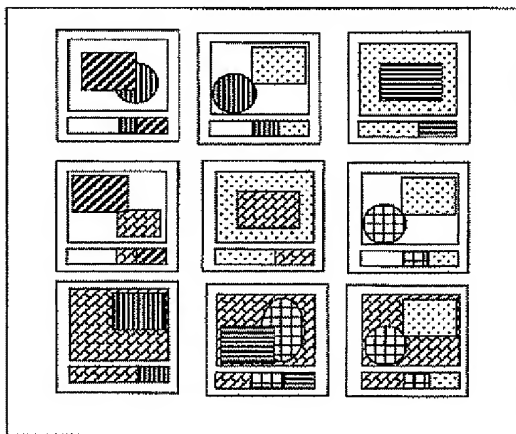
【図4】



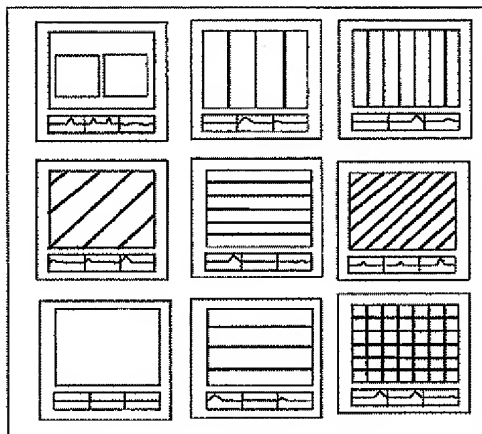
【図7】



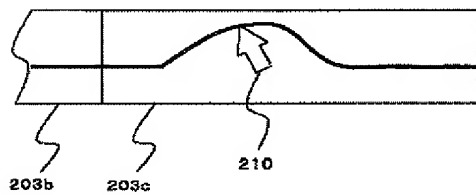
【図5】



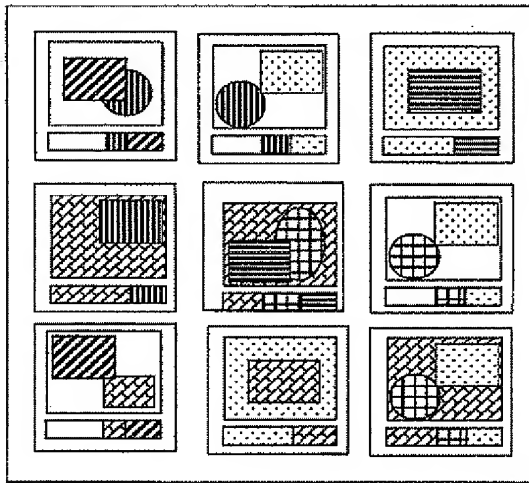
【図6】



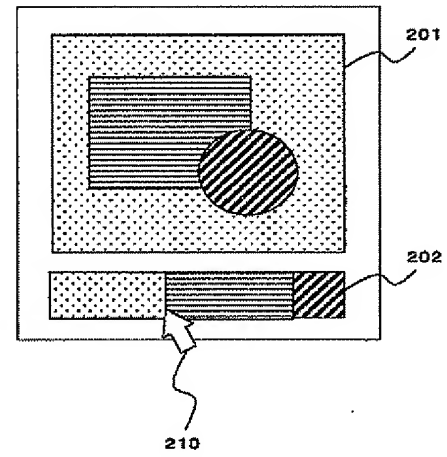
【図12】



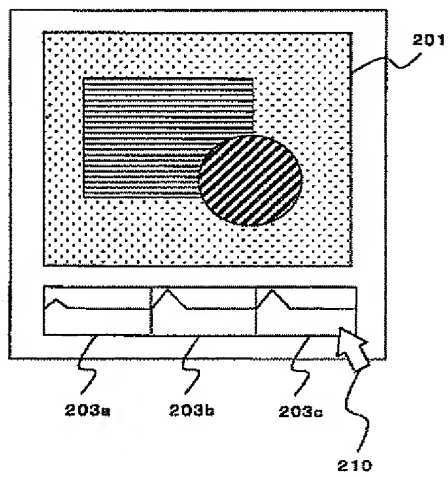
【図8】



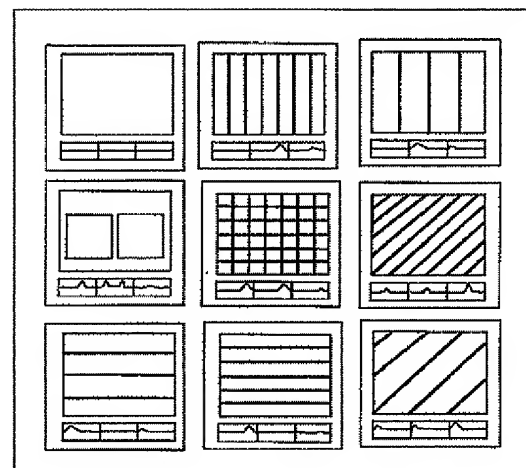
【図9】



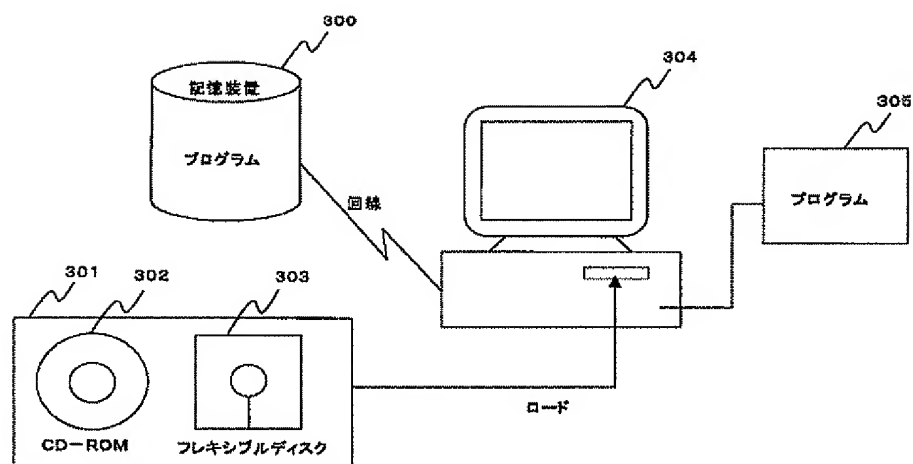
【図10】



【図11】



【図13】



フロントページの続き

(72)発明者 増本 大器
神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

Fターム(参考) 5B050 CA07 DA02 DA04 EA06 EA08
EA09 EA10 EA18 EA19 FA02
GA08
5B075 ND08 NK06 PP03 PP28 PQ40
PQ74
5L096 AA02 AA06 EA03 FA06 FA29
FA37 GA41 JA11 KA09